Spectro-Temporal Electrocardiogram Analysis for Noise-Robust Heart Rate and Heart Rate Variability Measurement

DIANA P. TOBÓN, SRINIVASAN JAYARAMAN, AND TIAGO H. FALK, (Senior Member, IEEE)

INRS-EMT, Université du Québec, Montreal, QC H5A-1K6, Canada

CORRESPONDING AUTHOR: TIAGO H. FALK (falk@emt.inrs.ca)

ABSTRACT The last few years has seen a proliferation of wearable electrocardiogram (ECG) devices in the market with applications in fitness tracking, patient monitoring, athletic performance assessment, stress and fatigue detection, and biometrics, to name a few. The majority of these applications rely on the computation of the heart rate (HR) and the so-called heart rate variability (HRV) index via time-, frequency-, or non-linear-domain approaches. Wearable/portable devices, however, are highly susceptible to artifacts, particularly those resultant from movement. These artifacts can hamper HR/HRV measurement, thus pose a serious threat to cardiac monitoring applications. While current solutions rely on ECG enhancement as a pre-processing step prior to HR/HRV calculation, existing artifact removal algorithms still perform poorly under extremely noisy scenarios. To overcome this limitation, we take an alternate approach and propose the use of a spectro-temporal ECG signal representation that we show separates cardiac components from artifacts. More specifically, by quantifying the rate-of-change of ECG spectral components over time, we show that heart rate estimates can be reliably obtained even in extremely noisy signals, thus bypassing the need for ECG enhancement. With such HR measurements in hands, we then propose a new noise-robust HRV index termed MD-HRV (modulation-domain HRV) computed as the standard deviation of the obtained HR values. Experiments with synthetic ECG signals corrupted at various different signal-to-noise levels, as well as recorded noisy signals show the proposed measure outperforming several HRV benchmark parameters computed post wavelet-based enhancement. These findings suggest that the proposed HR measures and derived MD-HRV metric are well-suited for ambulant cardiac monitoring applications, particularly those involving intense movement (e.g., elite athletic training).

INDEX TERMS Electrocardiogram, heart rate variability, modulation spectrum, telehealth, wearables.

I. INTRODUCTION

According to the American Heart Association, cardiovascular disease is one of the major causes of death globally with 17.3 million deaths per year. The number of deaths is expected to grow to more than 23.6 million by 2030 [1]. As such, electrocardiogram (ECG) monitoring technologies are burgeoning and portable/wearable devices have emerged, as well as numerous diagnostic and quantified self applications [2]. Such devices are capable of monitoring heart rate (HR), a measure of the number of contractions of the heart per minute (beats per minute, bpm), as well as the so-called heart rate variability (HRV) index, a measure of the variation in beat-to-beat intervals. HRV analysis has been used as a viable technique for non-invasive assessment of the automatic nervous system (ANS), both in healthy individuals and in patients with cardiac disorders (e.g., [3]–[8]). The ANS is comprised of the sympathetic and parasympathetic nervous systems, where the former prepares the body for action and maintains homeostasis, whilst the latter stimulates the body for relaxation. In response to such ANS activity, heartbeat intervals fluctuate causing a change in the variability of the heart rhythm. This variability has been shown to be related to cardiac autonomic function regulation [9], [10], as well as to stress, anxiety, diabetes, hypertension, fatigue, and drowsiness [11]–[15], fitness and sport assessment [16]–[20], as well as emotional states [21].

Over the years, several indices have been proposed to measure HRV, including linear (time domain, spectral domain) and non-linear methods (e.g., geometric method, entropy, and fractal dynamic methods). Time domain methods typically evaluate the variability in ECG peak-to-peak intervals (also known as 'RR' periods) and in the so-called normal-to-normal (NN) intervals between adjacent QRS complexes resulting from sinus node depolarizations [22]. As such, indices such as the standard deviation of the heart rate (i.e., inverse of the RR periods, termed sdHR), standard deviation of all NN intervals (SDNN), and standard deviation of the averages of NN intervals in all 5 min segments of the entire recording (SDANN) have been used. Frequency domain methods, in turn, apply either auto-regressive models or Welch periodogram analysis on the RR interval series in different frequency bands, such as ultra low frequency (ULF, $\leq$ 0.003 Hz), very low frequency (VLF, $0.003 - 0.04$ Hz), low frequency (LF, $0.04 - 0.15$ Hz), and high frequency (HF, $0.15 - 0.40$ Hz). LF and HF components have, in the past, been linked to sympathetic and parasympathetic nervous systems, respectively.

Non-linear methods, in turn, have been proposed and shown to better characterize the complex dynamics of the cardiac autonomic system [23]. As such, geometric methods have been used to convert the RR intervals into a geometric form (e.g triangle or ellipse), thus representing different classes of HRV [24], [25]. Other non-linear measures include indices of the signal "randomness" (entropy), with measures such as approximate, sample, multiscale, fuzzy, and fuzzy measure entropy [26], titration method [27], as well as statistical properties of fractals [28]–[30]. A complete description of such HRV indices is beyond the scope of this paper and the interested reader is referred to [22] and the references therein for more details.

Despite the method used to measure HRV, one commonality between all methods is their sensibility to ECG artifacts and to errors in the detection of instantaneous heart beats [31]–[33]. Movement artifacts are particularly troublesome, especially with wearable devices [34]. As such, robust peak detection algorithms have been proposed to better generate RR time series [35]. More recently, ECG enhancement (artifact removal) has also been explored to improve the ECG signal-to-noise ratio prior to peak detection. To this end, wavelet-based enhancement has shown to be powerful tool given its artifact removal capabilities and low computational power requirements [36]–[38]. Once computed, additional pre-processing steps, such as RR series detrending and inter-beat-interval resampling, are performed prior to HRV measurement. Notwithstanding, in extremely noisy scenarios, ECG enhancement and peak detection algorithm performances are limited [39], thus hampering HRV measurement and potentially leading to e.g., a rise in false alarms in automated health monitoring devices [40].

Here, we propose an alternate approach to noise-robust HR and HRV measurement based on a spectro-temporal ECG signal processing technique termed "modulation spectrum" processing. More specifically, we measure the rate-of-change of ECG short-term spectral magnitude components and show that cardiac components and noise components become separable in this new domain. Modulation spectrum processing has been shown in the past to accurately separate signal and noise components for speech enhancement [41], [42], auscultatory sound analysis [43], and, more recently, for automated ECG quality assessment [44]. Within this new domain, robust HR measurement in extremely noisy cases is possible without the need for a priori ECG enhancement nor time-domain peak detection. With such HR measures in hands, we propose a new HRV metric termed modulation-domain HRV (MD-HRV). This work builds on that of [45] and uses synthetic *and* recorded signals to validate the proposed measure against several time-, frequency, and non-linear domain HRV indices. To show the benefits of the proposed metric, comparisons are performed against these benchmark indices with and without a priori ECG enhancement; here, a wavelet-based enhancement method was used.

The remainder of this paper is organized as follows. Section II provides a description of the ECG modulation spectral representation, how it is used to measure HR and the proposed MD-HRV metric, benchmark HR/HRV metrics, and performance figures-of-merit. Experimental results are then presented in Section III and discussed in Section IV. Finally, conclusions and future research directions are drawn in Section V.

II. METHODS AND MATERIALS

This section presents the spectro-temporal ECG representation and describes the proposed HR calculation and the MD-HRV metric extraction approach. Further, a description of the synthetic and realistic datasets used, HRV metrics in time and frequency domains, as well as nonlinear measures are introduced. Finally, figures-of-merit to gauge the proposed HR calculation and MD-HRV performance are presented.

A. SPECTRO-TEMPORAL ECG REPRESENTATION

The processing steps for the computation of the spectro-temporal ECG representation are depicted in Figure 1. First, the ECG signal $x(n)$ (in our case, sampled at 256 Hz) is segmented into overlapping frames employing a 32-point sine window and 75% overlap to be then transformed to the frequency domain (spectrogram) via a 512-point fast Fourier transform (FFT). Then, the magnitude of the spectral components $\|X(f, m)\|$ (f is the conventional frequency in Hertz (Hz) and m is the frame index across time) is segmented using a 128-point sine window with 75% overlap to obtain the frequency-frequency representation $\mathcal{X}(f, f_{m,k})$ via a 512-point FFT. The frequency-frequency representation is known as modulation spectrogram, where $f_{m,k}$ is the modulation frequency in Hertz (Hz) and k is the frame index for the second FFT. The represented modulation spectrogram shows the rate-of-change of the different ECG spectral
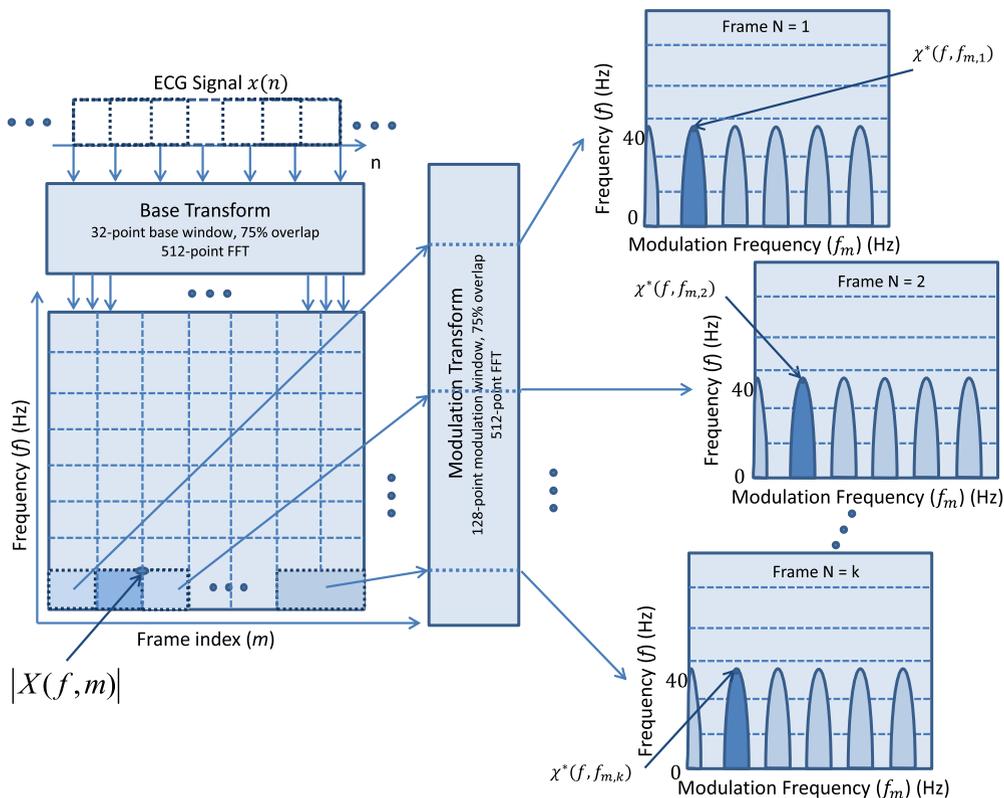
FIGURE 1. Signal processing steps involved in the spectro-temporal ECG representation. ECG signal, *x(n)*, is first segmented and transformed into the time-frequency representation via 512-point FFT. Modulation spectral magnitudes are then segmented and transformed via a second transform (512-point FFT) into a frequency-frequency representation. The right part of the figure shows modulation spectrograms frames from 1 to *k*.

components. Synthesized clean ECG data were used to optimize the base and modulation window sizes, as well as the overlap rates reported above for 256 Hz sampled ECG data, as detailed in [44]. This parameter configuration requires a minimum of 4 seconds of ECG data in order to calculate one modulation spectrogram.

Figure 2 (a) shows a typical modulation spectrogram of a synthetic clean ECG signal with 120 bpm. Figures 2 (b) and (c), in turn, depict its noisy counterparts with a signal-to-noise ratio (SNR) of 0 *dB* and −10 *dB*, respectively. As can be seen from Figure 2 (a), the first "lobe" centred at 2 Hz corresponds to the ECG heart rate, followed by several harmonics across the modulation frequency axis. While the first lobe energy decreases as the signal becomes noisier, it is still possible to detect it even at *SNR* = −10 *dB* (Figure 2 (c)), thus suggesting accurate HR measurement even at very noisy scenarios. The corresponding time domain waveforms can be seen in Figure 3. From the bottom plot, it can be seen that detecting heart beats and RR time series at an *SNR* = −10 *dB* could be very difficult, thus hampering conventional HRV measurement. The clear advantage of the modulation spectrum for heart rate detection in noisy ECGs is the main motivation for using this representation for HRV measurement, as described next.

B. MODULATION DOMAIN HR AND HRV MEASUREMENT

As outlined in Section II.A, the spectro-temporal ECG representation allows measuring heart rates even in extremely noisy scenarios, as depicted in Figure 2 (c). In order to compute the "instantaneous" heart rate (HR), the central frequency of the lobes, within each per-frame modulation spectrogram $\mathcal{X}^*(f, f_{m,k})$, has to be detected. This is represented by the shaded lobes in the modulation spectrograms of Figure 1. The central modulation frequency of the first lobe represents the HR (given in bpm when multiplied by 60), whilst the central frequency of the other lobes correspond to multiples (harmonics) of this HR value. In order to more accurately measure HR, the energy calculation for each modulation frequency bin is constrained to the $0 \leq f \leq 40$ range (where most of the ECG energy is concentrated) and the $0.8 \leq f_m \leq 3.3$ range, thus covering heart rates between 48 and 198 beats per minute. Such HR range is representative of the athletic and fitness applications explored herein, but could be easily expanded to values outside this range. As can be seen, with the proposed approach, heart rate measurement has shifted from the conventional time-domain amplitude peak detection approach, to a modulation frequency domain peak energy detection approach. Notwithstanding, since the latter relies on a signal representation in which cardiac

(a)

(b)

(c)
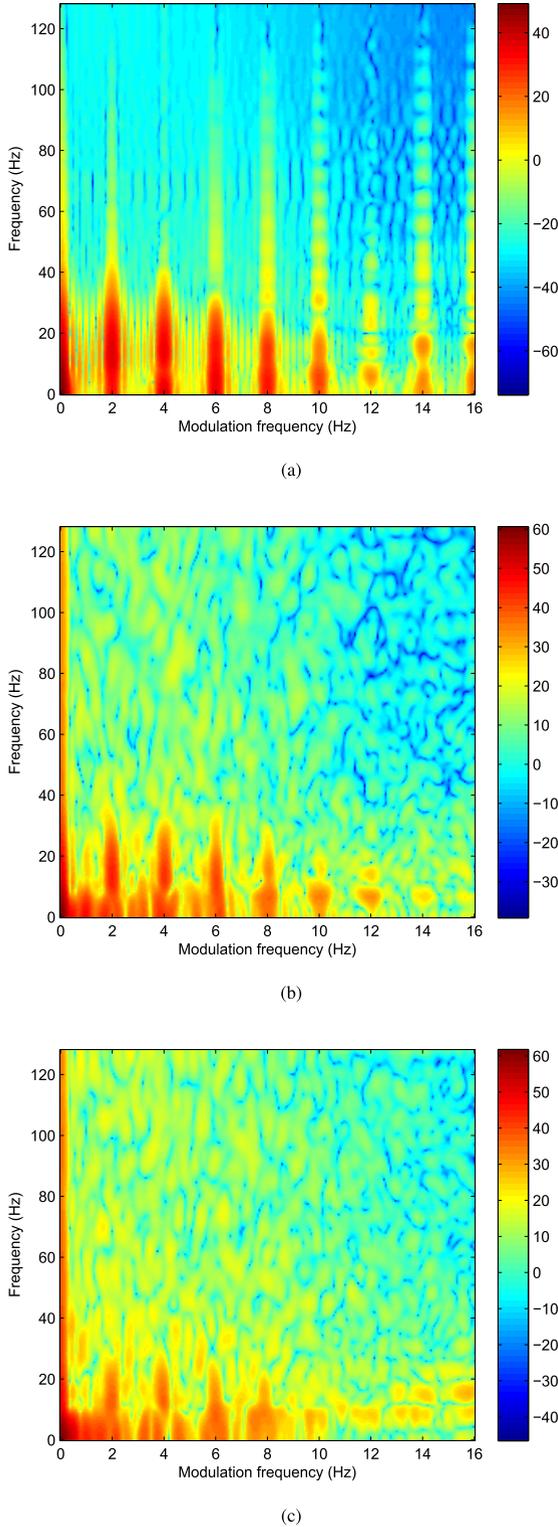
FIGURE 2. Modulation spectrograms of synthesized ECG with (a) clean 120 bpm and its noisy counterparts signals with an (b) SNR = 0 dB, and (c) SNR = −10 dB. The color bars show the range values in dB.

components and noise are separable, it becomes more robust to artifacts.

With the computed per-frame central modulation frequencies (i.e., "instantaneous" HRs), we propose the so-called modulation domain HRV metric (MD-HRV) as the standard deviation over all frames. When computed in beats per minute, MD-HRV is given as:

$$MD - HRV(\text{bpm}) = 60 \times \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}(\mathcal{X}^*_k - \bar{\mathcal{X}}^*)^2}, \quad (1)$$

where N is the total number of 'modulation frames', $\mathcal{X}^*_k$ is the central frequency (i.e., HR) at the kth modulation frame, and $\bar{\mathcal{X}}^*$ is the mean central frequency calculated over all per-frame modulation spectrograms. The multiplying factor of 60 is used to convert the MD-HRV metric from Hz to bpm. Here, modulation spectrograms are estimated every 5 seconds with 75% overlap. While alternate parameter configurations (e.g., base window, modulation window, overlap, and number of FFT points) can be used, the values reported herein were empirically found to be optimal for the task and sampling rate at hand [44]. For this particular configuration, a 5-second duration ECG signal will generate an RR series with five points. In the experiments herein, MD-HRV is computed over 10-minute ECG recordings, as detailed next.

C. DATASET 1: SYNTHETIC ECG
Synthetic ECG signals were generated using the 'ecgsyn' function in Matlab^{TM} available in Physionet [46]. The function uses a dynamic model to generate the ECG waveform described in [47] and allows configuring several parameters such as mean heart rate, sampling frequency, standard deviation of the heart rate, waveform morphology (i.e., P, Q, R, S, and T duration, timing, and amplitude), and low-frequency (LF) to high frequency (HF) ratio. Here, in total 700 signals of 10-minute duration sampled at 256 Hz were generated by randomly sampling the heart rate standard deviation between 1 and 10 bpm for heart rates ranging from 50 to 180 bpm. The heart rate standard deviation and heart rate ranges were chosen in order to cover cardiac diseases such as bradycardia, tachycardia, and arrhythmia as well as different activity levels (i.e., resting, walking, and running). Clean synthetic signals were corrupted by several artifacts at known SNR levels. Baseline wander noise, and muscle artifacts were taken from MIT-BIH Noise Stress Test Dataset [48]. Moreover, pink noise was added to model observation noise, as well as brownian noise to model electrode movement artifacts. Noisy signals were generated at 6 SNR levels of −10 dB, −8 dB, −5 dB, 0 dB, 5 dB, and 10 dB. Figure 3 depicts synthesized 5-second excerpts of a clean (top) and two noisy signals at SNR= −10 dB (bottom) and 0 dB (middle). Overall, 4200 noisy 10-minute signals (700 hours) were generated for HRV metric assessment.

D. DATASET 2: MIT-BIH ARRHYTHMIA DATASET
In order to test the proposed HRV metric in realistic signals, the MIT-BIH Arrhythmia database was also used [49]. Data sampled at 360 Hz were taken from 48 patients by the BIH Arrhythmia Laboratory, comprising two-channel ambulatory ECG recordings. Two cardiologists analyzed
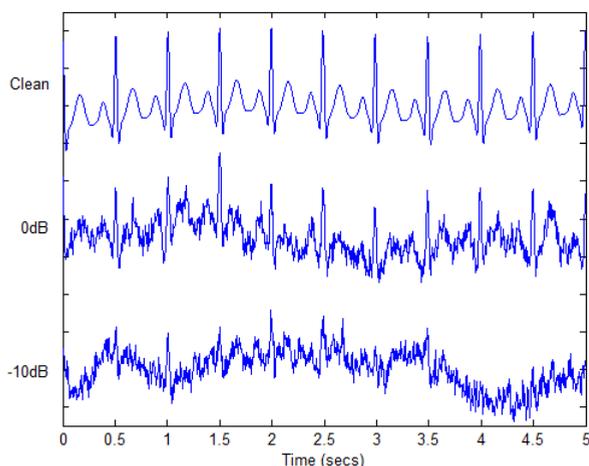
FIGURE 3. Five-second excerpts from the synthetic ECG signals with 120 bpm for clean and two noisy counterparts with an SNRs of 0 dB and −10 dB. Vertical axis is given in millivolts.

independently the data, where disagreements were resolved to obtain annotations for each beat. These annotation files allowed 'ground truth' HRV metrics to be calculated from the human-derived RR time series. Each recording has 30-minute duration. For the experiments described herein, the two-channel signals were dowsampled to 256 Hz and segmented into three 10-minute segments, thus similar to the synthetic data described above. Thus, a total of 288 10-minute segments were created providing an overall of 48 hours for testing. These segments were then grouped according to the heart rate (HR) and heart rate standard deviation (sdHR), similar to the synthetic ECG signals, thus allowing for easier comparison between the two databases. Hence, four groups were used: Group 1 (60 ± 10 bpm), Group 2 (80 ± 10 bpm), Group 3 (100 ± 10 bpm), and Group 4 (120 ± 10 bpm).

E. BENCHMARKS AND FIGURES-OF-MERIT

The popular Pan-Tompkins ECG peak detection algorithm was used to extract the RR time series [50] from the *clean* ECG signals. These RR series were used to calculate the 'ground truth' HR in the case of the synthetic ECG signals, and then input to the open-source HRVAS software package [51] to compute several benchmark HRV metrics based on time domain, frequency domain, and non-linear methods. For the recorded ECG dataset, human-derived RR times series were used for benchmark HR and HRV measurement. More specifically, time-domain HRV metrics include i) the widely-used SDNN, corresponding to the standard deviation of the normal-to-normal (NN) intervals, i.e., reflects cyclic components which cause variability in the ECG signal [52]; ii) SDANN, i.e., the standard deviation of 5-minute mean NN intervals, and iii) standard deviation of heart rate estimates (stdHR) computed after peak detection and given in bpm.

The frequency-domain metric, in turn, measures the total spectral power contained in frequency sub-bands such as very low frequency (VLF, 0-0.04 Hz), low frequency

(LF, 0.04-0.15 Hz), and high frequency (HF, 0.15-0.4 Hz). The total power was calculated using an auto-regressive Burg model order of 16 [53]. Lastly, non-linear Poincaré plots are used to quantify self-similarity assuming that each inter-beat interval (IBI) is influenced by the previous one [51]. Thus, an ellipse is fitted with a long axis representing the line of identity as well as a perpendicular axis to it. If the IBIs are longer than the previous ones, the points will be located above the line of identity and below for the opposite case. Consequently, the so-called SD2 parameter represents the standard deviation along the line of identity and the SD1 parameter the standard deviation over the perpendicular line. Hence, SD1 indicates the standard deviation of instantaneous beat-to-beat (i.e., short term variability) and SD2 of continuous or long-term variability [54].

As mentioned previously, a wavelet-based ECG enhancement algorithm [55] was used for artifact removal prior to peak detection and RR time series measurement (i.e, enhanced HR calculation). Such setup exemplifies the pre-processing pipeline typically used in existing state-of-the-art HRV and HR monitoring applications [35]. Several alternate enhancement algorithms were explored, such as Wiener filtering and ones based on empirical mode decomposition (EMD) [56]. Pilot experimentation [57] found the wavelet-based benchmark to provide the best results within the datasets used herein. Following insights from [58] and the pilot experiment, Daubechies-6 wavelets with 8 levels of decomposition, combined with soft thresholding and universal shrinkage were used [59]. To show the advantages of the proposed HRV metric, comparisons are made between MD-HRV computed directly from the *noisy* ECG signal and benchmark HRV metrics computed from the *noisy* and *enhanced* signals with RR time series computed using the Pan-Tompkins algorithm [50]. Moreover, the HR computed via time-domain RR time series analysis from the *noisy* and *enhanced* signals are compared to the HR values obtained via the proposed modulation spectrum domain method from the *noisy* data.

As for figures-of-merit, Pearson (ρ) correlation and HR error percentage were used to gauge the benefits of the proposed MD-HRV metric and HR measurement method, respectively. As mentioned above, due to the availability of the original clean synthetic signal and the annotation files for the arrythmia dataset, the proposed HRV metric and estimated HR values could be compared with 'ground truth' benchmark HRV and HR values. Pearson correlation shows the linear relationship between MD-HRV and the 'true' HRV metrics with higher values corresponding to improved metrics. HR error percentage, in turn, shows the precision of HR measurement in the modulation domain. Additionally, to evaluate the bias and to estimate an agreement interval between metrics, Bland-Altman plots were used [60]. A Bland-Altman plot depicts the differences (vertical axis) against the averages (horizontal axis), hence, are used to verify the level of agreement between the 'true' metrics and MD-HRV [61]. Typically, Bland-Altman plots with a more

TABLE 1. Performance comparison (ρ) of HRV metrics for noisy signals, wavelet enhanced signals, and the proposed MD-HRV metric for Dataset 1 at different SNR levels.

SDNN			
Input SNR (dB)	Noisy	Enhanced	MD-HRV
−10	0.221 ± 0.375	0.439 ± 0.368	0.781 ± 0.055
−8	0.505 ± 0.474	0.615 ± 0.383	0.924 ± 0.069
−5	0.896 ± 0.182	0.852 ± 0.281	0.969 ± 0.045
0	0.912 ± 0.215	0.965 ± 0.091	0.979 ± 0.030
5	1.000 ± 0.001	0.976 ± 0.062	0.988 ± 0.009
10	1.000 ± 0.000	1.000 ± 0.000	0.990 ± 0.048

SDANN			
Input SNR (dB)	Noisy	Enhanced	MD-HRV
−10	0.271 ± 0.382	0.443 ± 0.353	0.751 ± 0.074
−8	0.541 ± 0.464	0.650 ± 0.353	0.915 ± 0.071
−5	0.920 ± 0.141	0.859 ± 0.266	0.962 ± 0.049
0	0.944 ± 0.119	0.970 ± 0.061	0.970 ± 0.040
5	0.994 ± 0.004	0.985 ± 0.027	0.980 ± 0.019
10	0.996 ± 0.003	0.998 ± 0.001	0.982 ± 0.016

sdHR			
Input SNR (dB)	Noisy	Enhanced	MD-HRV
−10	0.225 ± 0.115	0.285 ± 0.346	0.762 ± 0.079
−8	0.537 ± 0.162	0.494 ± 0.314	0.924 ± 0.066
−5	0.883 ± 0.099	0.836 ± 0.213	0.969 ± 0.045
0	0.989 ± 0.025	0.991 ± 0.015	0.980 ± 0.029
5	0.999 ± 0.001	0.996 ± 0.006	0.990 ± 0.010
10	0.999 ± 0.000	0.999 ± 0.001	0.991 ± 0.007

SD1			
Input SNR (dB)	Noisy	Enhanced	MD-HRV
−10	0.156 ± 0.298	0.407 ± 0.359	0.757 ± 0.076
−8	0.443 ± 0.270	0.562 ± 0.379	0.916 ± 0.070
−5	0.821 ± 0.290	0.794 ± 0.289	0.963 ± 0.041
0	0.802 ± 0.467	0.906 ± 0.218	0.973 ± 0.027
5	0.969 ± 0.061	0.887 ± 0.295	0.981 ± 0.008
10	0.993 ± 0.003	0.999 ± 0.001	0.983 ± 0.005

SD2			
Input SNR (dB)	Noisy	Enhanced	MD-HRV
−10	0.267 ± 0.415	0.450 ± 0.367	0.759 ± 0.077
−8	0.528 ± 0.521	0.638 ± 0.381	0.921 ± 0.066
−5	0.917 ± 0.126	0.864 ± 0.278	0.963 ± 0.052
0	0.952 ± 0.110	0.980 ± 0.050	0.974 ± 0.036
5	0.998 ± 0.004	0.991 ± 0.024	0.982 ± 0.020
10	0.998 ± 0.004	1.000 ± 0.000	0.984 ± 0.017

Total power (AR)			
Input SNR (dB)	Noisy	Enhanced	MD-HRV
−10	0.194 ± 0.265	0.346 ± 0.322	0.711 ± 0.067
−8	0.423 ± 0.355	0.524 ± 0.408	0.871 ± 0.049
−5	0.803 ± 0.273	0.760 ± 0.317	0.908 ± 0.042
0	0.791 ± 0.329	0.889 ± 0.259	0.918 ± 0.029
5	0.921 ± 0.018	0.922 ± 0.188	0.926 ± 0.027
10	0.924 ± 0.023	0.998 ± 0.002	0.928 ± 0.028

TABLE 2. Performance comparison (ρ) of HRV metrics for noisy signals, wavelet enhanced signals, and the proposed MD-HRV metric for Dataset 2.

HRV metrics	Noisy	Enhanced	MD-HRV
SDNN	0.632 ± 0.316	0.617 ± 0.233	0.893 ± 0.046
SDANN	0.687 ± 0.154	0.641 ± 0.026	0.899 ± 0.022
sdHR	0.608 ± 0.288	0.616 ± 0.318	0.902 ± 0.044
SD1	0.675 ± 0.135	0.632 ± 0.141	0.908 ± 0.017
SD2	0.611 ± 0.343	0.601 ± 0.254	0.879 ± 0.059
Total Power (AR)	0.452 ± 0.202	0.281 ± 0.149	0.893 ± 0.060

TABLE 3. Performance comparison (HR error percentage) of HR measurement for noisy signals, wavelet enhanced signals, and the proposed approach for Dataset 1.

Input SNR (dB)	Noisy	Enhanced	Proposed
−10	58.8 ± 73.3	7.7 ± 9.9	1.7 ± 2.6
−8	40.9 ± 61.7	2.3 ± 3.1	0.9 ± 0.8
−5	6.5 ± 21.9	0.7 ± 0.9	0.5 ± 0.6
0	0.1 ± 0.2	0.2 ± 0.4	0.3 ± 0.5

ECG signals, their enhanced wavelet counterparts, and the proposed MD-HRV metric for datasets 1 and 2, respectively. The Pearson (ρ) correlation was obtained between the 'true' benchmark HRV metrics for clean signals (i.e., SDNN, SDANN, sdHR, total power, SD1, and SD2) and HRV metrics for noisy and enhanced signals, as well as between the 'true' metrics and the proposed MD-HRV. The results reported in Table 2 correspond to the average over the four groups, as described in Section II-D. Table 3, in turn, reports the HR estimation error (expressed as percentage of the true HR) obtained via time-domain RR series analysis from the noisy and enhanced ECG signals, as well as with the proposed method. These results correspond to the average over all per-frame modulation spectrograms (computed every 5 seconds herein). Focus is placed on Dataset 1 to explore the effects as a function of very low SNRs below 0 dB.

Figures 4 (a)-(c) further depict the scatterplots (and linear fit curves) of the sdHR metric computed from the noisy (subplot a) and enhanced signals (subplot b), as well as the MD-HRV metric computed from the noisy signals (subplot c). The sdHR metric is chosen as it closely resembles the processing performed by the MD-HRV metric, but with RR. To avoid cluttered plots, only 100 noisy signals at 100 bpm and SNR=−10 dB are shown, thus representing scenarios similar to high levels of exercise and movement. Furthermore, Figures 5 (a)-(c) show the Bland-Altman plots for the three cases mentioned above, respectively.

Lastly, Figures 6 and 7 show the scatterplots and Bland-Altman plots for Group 2 (80 ± 10 bpm) in dataset 2, respectively. As above, the sdHR benchmark metric computed from the noisy signal, the enhanced signal, as well as the proposed MD-HRV metric computed from the noisy signal are shown in the plots. It is clear from the plots that the proposed MD-HRV metric results in higher correlations, tighter regression confidence intervals, and more compact Bland-Altman plots around zero (vertical axis) than the sdHR benchmark metric in both cases with and without wavelet enhancement.

compact point distribution around zero (vertical axis) represent metrics in better agreement. Lastly, linear regression analysis is performed to show the linear fitting with confidence intervals at 95%; smaller confidence intervals represent more reliable metrics.

III. RESULTS
Tables 1 and 2 show the performance comparison for the different HRV metrics calculated from the noisy

FIGURE 4. Scatterplots for noisy ECG signals with 100 bpm and SNR = −10 dB between (a) 'true' sdHR and noisy sdHR, (b) 'true' sdHR and wavelet enhanced sdHR, and (c) 'true' sdHR and proposed MD-HRV.

FIGURE 5. Bland-Altman plots for noisy ECG signals with 100 bpm and SNR = −10 dB between (a) 'true' sdHR and noisy sdHR, (b) 'true' sdHR and wavelet enhanced sdHR, and (c) 'true' sdHR and proposed MD-HRV.

IV. DISCUSSION

In this paper, a noise-robust method to estimate HR and a new HRV metric were proposed based on a spectro-temporal ECG representation termed 'modulation spectrum.' The goal of the metric was to reliably measure HR/HRV in extremely noisy settings directly from the noisy ECG signal, without the need for a priori ECG enhancement nor time-domain peak detection. The proposed HRV metric was compared to six benchmark HRV indices computed using time-domain, frequency-domain and non-linear methods reported in the literature, with and without ECG enhancement and with RR time series computed using the popular Pan-Tompkins algorithm. The proposed HR measurement method, in turn, was compared to conventional RR time
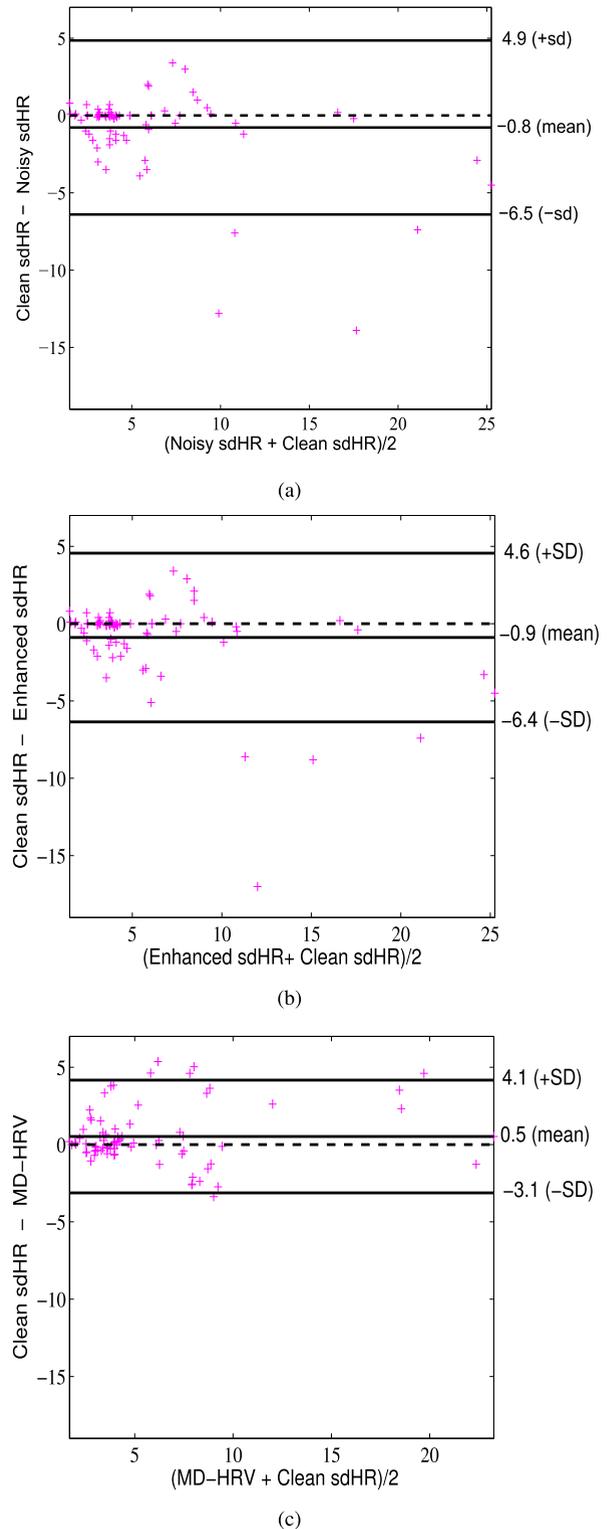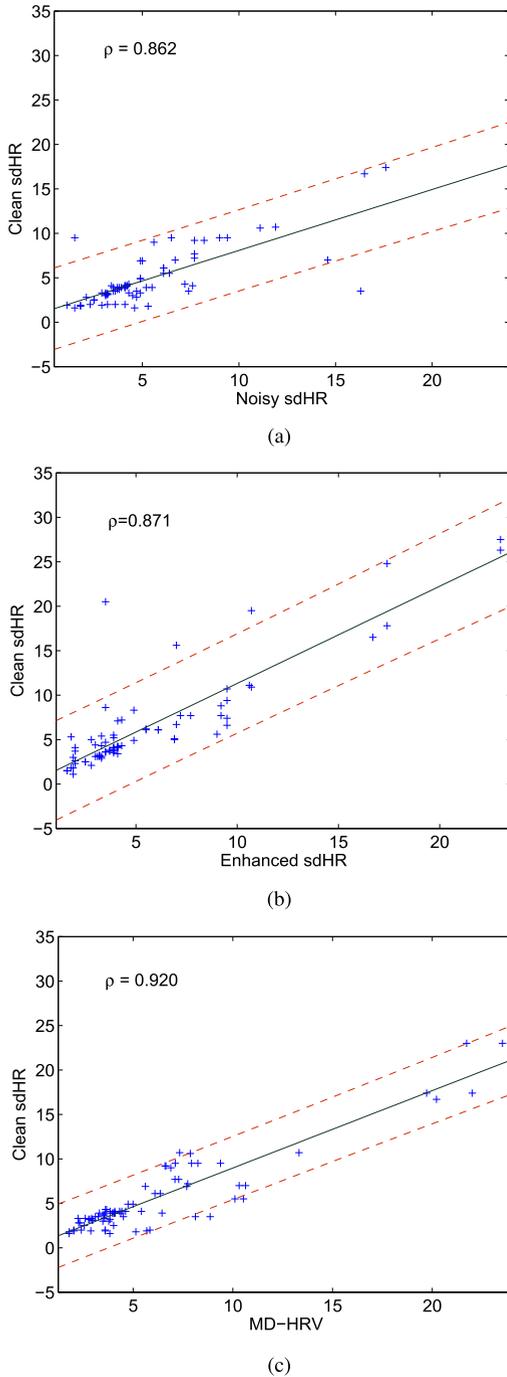
FIGURE 6. Scatterplots for Group 2 in Dataset 2 for (a) Noisy sdHR and 'true' sdHR, (b) Wavelet enhanced sdHR and 'true' sdHR, and (c) Proposed MD-HRV and 'true' sdHR.

FIGURE 7. Bland-Altman plots for Group 2 in Dataset 2 for (a) Noisy sdHR and 'true' sdHR, (b) Wavelet enhanced sdHR and 'true' sdHR, and (c) Proposed MD-HRV and 'true' sdHR.

series analysis computed from the noisy and enhanced signals.

From Table 1, it can be seen that benchmark HRV metrics computed from noisy synthetic signals are generally well correlated with the true HRV values computed from the clean ECG signals for SNR levels greater than 0 dB. For extremely low SNR levels below -5 dB, however,

typical of those observed with wearables during intense exercise [44], HRV measurement accuracy degrades quickly. Wavelet-based enhancement, in turn, was shown to be useful for most metrics, particularly for SNRs below -8 dB.

The proposed MD-HRV metric, on the other hand, computed only from the *noisy* ECG signal, showed stable correlation values generally greater than 0.9 up to an SNR= $-8dB$, with a drop to around $0.71 - 0.76$ at an SNR= -10 dB. In this extremely noisy scenario, it was difficult to accurately detect the main lobe in the modulation spectrogram (corresponding to the heart rate, see Section II-A), thus compromising HRV measurement. Notwithstanding, it can be seen that with the SDNN metric, while wavelet enhancement improved ρ by about 99% relative to the noisy case, the proposed metric outperformed the noisy case by 253%. Similar gains were seen across all of the benchmark metrics, including SDANN (177% compared to 63% post enhancement), sdHR (239% compared to 27% post enhancement), SD1 and SD2 (385% and 184% compared to 161% and 68% post enhancement, respectively), and total power (266% compared to 78% post enhancement). Overall, the correlation values obtained with the MD-HRV metric at an SNR= -10 dB are inline with those obtained with the benchmark metrics *post* enhancement at -8 dB $\leq$ SNR ≤ -5 dB. Closer inspection showed that the total power HRV benchmark metric resulted in the lowest correlation with MD-HRV across all tested SNR levels. Interestingly, this benchmark metric showed to be the most sensitive to artifacts, and resulted in the lowest correlation values of all benchmark metrics even for SNR= 10 dB.

These results are further validated from the scatter and Bland-Altman plots shown in Figures 4 and 5, respectively, for the proposed and sdHR (pre and post enhancement) metrics; plots correspond to 100 randomly-selected ECG signals at 100 bpm and an SNR= -10 dB. The Bland-Altman plots, for example, show that the discrepancies between the true sdHR (computed from the clean signals) and the sdHR computed from the wavelet enhanced signals had a mean difference of 11.3 ± 8.1, whereas this difference was of 2.2 ± 3.1 for the proposed MD-HRV metric computed from the noisy signals. Overall, results from Table 1 and Figures 4 and 5 suggest that the proposed metric could be useful in extreme conditions where wearable ECG signals are highly contaminated with motion artifacts, such as during peak performance training.

Table 2, in turn, reported results obtained from ambulatory recorded ECG data. Relative to Table 1, the correlation values suggest a fair amount of noise is present in the recordings. Interestingly, wavelet based enhancement did not result in performance improvements in terms of ρ correlation for most of the benchmark metrics, thus highlighting the limitations of existing enhancement algorithms for data recorded in real-world settings. Notwithstanding, the proposed metric achieved reliable results in terms of ρ correlation without the need for a-priori enhancement; overall, ρ values stayed between $0.8 - 0.9$ for all six benchmark metrics. Overall, an average gain in ρ of approximately 49% was observed with the proposed MD-HRV metric over the wavelet-enhanced case. As in the case for synthetic ECG signals, the frequency-domain HR benchmark metric showed to be the most sensitive to artifacts. In fact, for Dataset 2,

wavelet enhancement deteriorated benchmark HRV measurement performance, likely due to the introduction of additional unwanted artifacts post enhancement. As with the synthetic signals, the scatter and Bland-Altman plots in Figures 6 and 7 show smaller confidence intervals and tighter distributions around the null bias, respectively. Overall, the proposed MD-HRV metric achieved a mean difference of 0.5 ± 3.6 with the true sdHR computed from the original heart rate labels, thus comparing favourably to the -0.9 ± 5.5 obtained with the wavelet enhanced benchmark metric.

Table 3 reports the HR estimation error (as a percentage of the true HR) obtained from Dataset 1, averaged over all per-frame modulation spectrograms for very low SNR levels. As can be seen, noisy ECG severely degrades heart rate estimation accuracy and errors greater than 40% can be seen for SNRs ≤ -8 dB. Wavelet enhancement, in turn, significantly improves HR measurement, but still results in a 7% HR estimation error. The error deviation is also high, with a standard deviation of 9.9% in the case of SNR = -10 dB. The proposed method, on the other hand, maintained errors below 2% and error deviation at 2.6%, thus suggesting more stable and precise measurement. Though not listed in the table, the same analysis was performed on Dataset 2 and the HR error (percentage) across all the 4 groups was 7.7 ± 22.4 for the noisy data, 7.9 ± 11.6 for wavelet enhanced data, and 3.0 ± 4.3 with the proposed technique. These findings corroborate those of Table 2 where the impact of enhancement was not as pronounced for real-world recorded ECGs. It becomes clear from this analysis, however, that wavelet enhancement reduces the variability of the HR estimates, thus provides some advantages over processing the noisy signals. Overall, the proposed technique results in the lowest error and error variability, thus suggesting a better suited method for realistic scenarios.

In terms of computational complexity, the computation time to calculate the MD-HRV metric for a 10-minute duration ECG signal was approximately 1.76 seconds. By processing a number of ECG signals of varying duration ranging from 1 minute to 10 minutes, it was observed that processing time increased linearly with the number of time samples. For the benchmark system, processing time for wavelet enhancement, Pan-Tompkins peak detection and HRV measurement ranged from 1.2-2.4 seconds for the SDNN and AR HRV metrics, respectively. It is important to emphasize that the MD-HRV Matlab codes developed herein have not been optimized for speed, thus faster processing times could still be achieved. Overall, these aforementioned findings suggest that the MD-HRV metric is indeed an improved noise-robust surrogate to existing HRV metrics with a small computational footprint. This can have important benefits for athletic and fitness applications in which SDNN, SDANN, SD1, SD2 and total power metrics have been widely used. For example, they were used to monitor the cardiac autonomic activity in free diving athletes in [17], for healthy elite/master athletes in [62] and for teenager athletes in [19], as well as in judo training in [18].

V. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

This paper has proposed a new method of estimating heart rate and heart rate variability in extremely noisy settings. The proposed metrics are based on the spectro-temporal representation of the ECG signal, commonly termed "modulation spectral" representation. The proposed HR and MD-HRV (modulation domain HRV) metrics were extensively tested using both synthetic and recorded noisy ECG signals. MD-HRV was compared to six benchmark HRV metrics computed from the noisy, as well as enhanced ECG signals, which had artifacts removed using a state-of-the-art wavelet-based algorithm. HR measurement performance, in turn, was gauged using HR measures computed using time-domain peak detection algorithms and RR series analysis. HR and HRV values computed using the proposed methods from the *noisy* signals outperformed those obtained from the benchmarks using the *enhanced* signals whilst requiring comparable computational processing time. Overall, these findings suggest that the proposed HR/HRV metrics can play key roles applications involving wearable devices under extreme movement conditions (e.g., high performance athletics). While the synthetic ECGs (Dataset 1) used herein were generated by manipulating parameters to simulate different cardiac diseases and the recorded data (Dataset 2) contained a few instances of arrhythmia, further research is still needed to validate the use of the proposed metrics within more clinically-driven applications (e.g., sleep apnea and congestive heart failure). Moreover, given the burgeoning of wearable heart rate devices based on photoplethysmograms, future research will explore the use of MD-HRV computed from alternate modalities beyond ECGs.

REFERENCES

[1] D. Mozaffarian *et al.*, "Heart disease and stroke statistics-2016 update: A report from the American Heart Association," *Circulation*, vol. 133, no. 4, p. e38, 2016.

[2] D. P. Tobón, T. H. Falk, and M. Maier, "Context awareness in WBANs: A survey on medical and non-medical applications," *IEEE Wireless Commun.*, vol. 20, no. 4, pp. 30–37, Aug. 2013.

[3] B. M. Appelhans and L. J. Luecken, "Heart rate variability as an index of regulated emotional responding," *Rev. Gen. Psychol.*, vol. 10, no. 3, pp. 229–240, 2006.

[4] S. Akselrod, D. Gordon, F. A. Hubel, D. C. Shannon, A. C. Barger, and R. J. Cohen, "Power spectrum analysis of heart rate fluctuation: A quantitative probe of beat-to-beat cardiovascular control," *Science*, vol. 213, pp. 220–222, Jul. 1981.

[5] A. Malliani, M. Paganiand F. Lombardi, and S. Cerutti, "Cardiovascular neural regulation explored in the frequency domain," *Circulation*, vol. 84, no. 2, pp. 482–492, 1991.

[6] A. Canim and L. Fei, "Clinical significance of heart rate variability," in *Noninvasive Electrocardiology: Clinical Aspects of Holter Monitoring*, A. J. Moss and S. Stem, Eds. London, U.K.: W.B. Saunders, 1996, pp. 225–248.

[7] A. Malliani, "The pattern of sympathovagal balance explored in the frequency domain," *Physiology*, vol. 14, no. 3, pp. 111–117, 1999.

[8] S. Priori *et al.*, "Task force on sudden cardiac death, european society of cardiology: Summary of recommendations," *Europace*, vol. 4, no. 1, pp. 3–18, 2002.

[9] B. Ram *et al.*, "Circadian heart rate and blood pressure variability considered for research and patient care," *Int. J. Cardiol.*, vol. 87, no. 1, pp. 9–28, 2003.

[10] A. Bilan, A. Witczak, R. Palusiński, W. Myśliński, and J. Hanzlik, "Circadian rhythm of spectral indices of heart rate variability in healthy subjects," *J. Electrocardiol.*, vol. 38, no. 3, pp. 239–243, 2005.

[11] B. Aysin, J. Colombo, and E. Aysin, "Comparison of HRV analysis methods during orthostatic challenge: HRV with respiration or without?" in *Proc. 29th Annu. Int. Conf. Eng. Med. Biol. Soc. (EMBS)*, Aug. 2007, pp. 5047–5050.

[12] I. Gritti, S. Defendi, C. Mauri, G. Banfi, P. Duca, and G. S. Roi, "Heart rate variability, standard of measurement, physiological interpretation and clinical use in mountain marathon runners during sleep and after acclimatization at 3480 m," *J. Behav. Brain Sci.*, vol. 3, no. 1, pp. 26–48, 2013.

[13] W. S. Liew, M. Seera, C. K. Loo, E. Lim, and N. Kubota, "Classifying stress from heart rate variability using salivary biomarkers as reference," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 10, pp. 2035–2046, Oct. 2016.

[14] G. Valenza, L. Citi, C. Gentili, A. Lanata, E. P. Scilingo, and R. Barbieri, "Characterization of depressive states in bipolar patients using wearable textile technology and instantaneous heart rate variability assessment," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 1, pp. 263–274, Jan. 2015.

[15] R. Lacuesta, L. Garcia, I. García-Magariño, and J. Lloret, "System to recommend the best place to live based on wellness state of the user employing the heart rate variability," *IEEE Access*, vol. 5, pp. 10594–10604, 2017.

[16] G. Sandercock, L. Hodges, P. Bromley, and D. Brodie, "Differences in autonomic control during exercise in fit and unfit subjects measured by heart rate variability," *Med. Sci. Sports Exercise*, vol. 36, p. S128, May 2004.

[17] V. Christoforidi, N. Koutlianos, P. Deligiannis, E. Kouidi, and A. Deligiannis, "Heart rate variability in free diving athletes," *Clin. Physiol. Funct. Imag.*, vol. 32, no. 2, pp. 162–166, 2012.

[18] P. S. Araújo *et al.*, "Cardiac autonomic modulation in judo athletes: Evaluation by linear and non-linear method," *Sport Sci. Health*, vol. 12, no. 1, pp. 125–130, 2016.

[19] V. K. Sharma, S. K. Subramanian, V. Arunachalam, and R. Rajendran, "Heart rate variability in adolescents—Normative data stratified by sex and physical activity," *J. Clin. Diagnostic Res.*, vol. 9, no. 10, pp. CC08–CC13, 2015.

[20] D. Jarchi and A. J. Casson, "Towards photoplethysmography-based estimation of instantaneous heart rate during physical activity," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 9, pp. 2042–2053, Sep. 2017.

[21] M. Nardelli, G. Valenza, A. Greco, A. Lanata, and E. P. Scilingo, "Recognizing emotions induced by affective sounds through heart rate variability," *IEEE Trans. Affect. Comput.*, vol. 6, no. 4, pp. 385–394, Oct. 2015.

[22] A. J. Camm *et al.*, "Heart rate variability: Standards of measurement, physiological interpretation and clinical use. Task force of the european society of cardiology and the north american society of pacing and electrophysiology, " *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996.

[23] K. T. Park and S. H. Yi, "Accessing physiological complexity of HRV by using threshold-dependent symbolic entropy," *J. Korean Phys. Soc.*, vol. 44, no. 3, pp. 569–576, Mar. 2004.

[24] C. K. Karmakar, A. H. Khandoker, H. F. Jelinek, and M. Palaniswami, "Risk stratification of cardiac autonomic neuropathy based on multi-lag tone–entropy," *Med. Biol. Eng. Comput.*, vol. 51, no. 5, pp. 537–546, 2013.

[25] N. Karim, J. A. Hasan, and S. S. Ali, "Heart rate variability—A review," *J. Basic Appl. Sci.*, vol. 7, no. 1, pp. 71–77, 2011.

[26] C. Liu *et al.*, "Analysis of heart rate variability using fuzzy measure entropy," *Comput. Biol. Med.*, vol. 43, no. 2, pp. 100–108, 2013.

[27] C.-S. Poon and M. Barahona, "Titration of chaos with added noise," *Proc. Nat. Acad. Sci. USA*, vol. 98, no. 13, pp. 7107–7112, Jun. 2001.

[28] J. M. Tapanainen *et al.*, "Fractal analysis of heart rate variability and mortality after an acute myocardial infarction," *Amer. J. Cardiol.*, vol. 90, no. 4, pp. 347–352, 2002.

[29] T. H. Mäkikallio *et al.*, "Fractal analysis and time- and frequency-domain measures of heart rate variability as predictors of mortality in patients with heart failure," *Amer. J. Cardiol.*, vol. 87, no. 2, pp. 178–182, 2001.

[30] R. A. U, P. S. Bhat, N. Kannathal, A. Rao, and C. M. Lim, "Analysis of cardiac health using fractal dimension and wavelet transformation," *ITBM-RBM*, vol. 26, no. 2, pp. 133–139, 2005.

[31] M. Malik *et al.*, "Influence of the noise and artefact in automatically analysed long term electrocardiograms on different methods for time-domain measurement of heart rate variability," in *Proc. Comput. Cardiol.*, Sep. 1991, pp. 269–272.

[32] B. Oke, B. Dietmar, M. Lutz, S. Jens, and H. Robert, "Impact of artefacts during routine ECG recording for analysis of heart rate variability: 3AP1-6," *Eur. J. Anaesthesiol.*, vol. 25, p. 26, May/Jun. 2008.

[33] M. Singh, B. Singh, and A. Rajput, "Impact of Missing RR-interval on Non-Linear HRV Parameters," *Int. J. Comput. Appl.*, vol. 124, no. 11, pp. 13–18, 2015.

[34] M. Elgendi, B. Eskofier, S. Dokos, and D. Abbott, "Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems," *PLoS ONE*, vol. 9, no. 1, p. e84018, 2014.

[35] N. R. Lang *et al.*, "Filter and processing method to improve R-peak detection for ECG data with motion artefacts from wearable systems," in *Proc. Comput. Cardiol. Conf. (CinC)*, Sep. 2015, pp. 917–920.

[36] H. He, Z. Wang, and Y. Tan, "Noise reduction of ECG signals through genetic optimized wavelet threshold filtering," in *Proc. IEEE Int. Conf. Comput. Intell. Virtual Environ. Meas. Syst. Appl. (CIVEMSA)*, Jun. 2015, pp. 1–6.

[37] B. N. Singh and A. K. Tiwari, "Optimal selection of wavelet basis function applied to ECG signal denoising," *Digit. Signal Process.*, vol. 16, no. 3, pp. 275–287, 2006.

[38] P. Karthikeyan, M. Murugappan, and S. Yaacob, "ECG signal denoising using wavelet thresholding techniques in human stress assessment," *Int. J. Elect. Eng. Inform.*, vol. 4, no. 2, p. 306, 2012.

[39] D. P. Tobón and T. H. Falk, "Adaptive modulation spectral filtering for improved electrocardiogram quality enhancement," in *Proc. Comput. Cardiol. (CinC)*, Sep. 2016, pp. 441–444.

[40] C. Tsimenidis and A. Murray, "False alarms during patient monitoring in clinical intensive care units are highly related to poor quality of the monitored electrocardiogram signals," *Physiol. Meas.*, vol. 37, no. 8, p. 1383, 2016.

[41] J.-W. Hung, W.-H. Tu, and C.-C. Lai, "Improved modulation spectrum enhancement methods for robust speech recognition," *Signal Process.*, vol. 92, no. 11, pp. 2791–2814, 2012.

[42] T. H. Falk, S. Stadler, W. B. Kleijn, and W. Y. Chan, "Noise suppression based on extending a speech-dominated modulation band," in *Proc. INTERSPEECH*, 2007, pp. 970–973.

[43] T. H. Falk, E. Sejdic, T. Chau, and W.-Y. Chan, "Spectro-temporal analysis of auscultatory sounds," in *Biomedical Engineering*. Rijeka, Croatia: InTech, 2010.

[44] D. P. Tobón, T. H. Falk, and M. Maier, "MS-QI: A modulation spectrum-based ECG quality index for telehealth applications," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 8, pp. 1613–1622, Aug. 2016.

[45] D. P. Tobón, S. Jayaraman, and T. H. Falk, "Improved heart rate variability measurement based on modulation spectral processing of noisy electrocardiogram signals," in *Proc. IEEE 14th Int. Conf. Wearable Implant. Body Sensor Netw. (BSN)*, May 2017, pp. 51–54.

[46] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[47] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 289–294, Mar. 2003.

[48] G. B. Moody, W. K. Muldrow, and R. G. Mark, "A noise stress test for arrhythmia detectors," *Comput. Cardiol.*, vol. 11, no. 3, pp. 381–384, 1984.

[49] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, May/Jun. 2001.

[50] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985.

[51] J. T. Ramshur, "Design, evaluation, and application of heart rate variability analysis software (HRVAS)," M.S. thesis, Dept. Biomed. Eng., Univ. Memphis, Memphis, TN, USA, 2010.

[52] M. Moser *et al.*, "Heart rate variability as a prognostic tool in cardiology. A contribution to the problem from a theoretical point of view," *Circulation*, vol. 90, no. 2, pp. 1078–1082, 1994.

[53] A. Boardman, F. S. Schlindwein, and A. P. Rocha, "A study on the optimum order of autoregressive models for heart rate variability," *Physiol. Meas.*, vol. 23, no. 2, p. 325, 2002.

[54] P. W. Kamen and A. M. Tonkin, "Application of the Poincaré plot to heart rate variability: A new measure of functional status in heart failure," *Int. Med. J.*, vol. 25, no. 1, pp. 18–26, 1995.

[55] P. B. Patil and M. S. Chavan, "A wavelet based method for denoising of biomedical signal," in *Proc. Int. Conf. Pattern Recognit., Inform. Med. Eng. (PRIME)*, Mar. 2012, pp. 278–283.

[56] M. A. Kabir and C. Shahnaz, "Denoising of ECG signals based on noise reduction algorithms in EMD and wavelet domains," *Biomed. Signal Process. Control*, vol. 7, no. 5, pp. 481–489, 2012.

[57] D. P. Tobón and T. Falk, "Adaptive spectro-temporal filtering for electrocardiogram signal enhancement," *IEEE J. Biomed. Health Inform.*, to be published, doi: 10.1109/JBHI.2016.2638120.

[58] S. Pal and M. Mitra, "Detection of ECG characteristic points using multiresolution wavelet analysis based selective coefficient method," *Measurement*, vol. 43, no. 2, pp. 255–261, 2010.

[59] G. Georgieva-Tsaneva and K. Tcheshmedjiev, "Denoising of electrocardiogram data with methods of wavelet transform," in *Proc. Int. Conf. Comput. Syst. Technol.*, 2013, pp. 9–16.

[60] D. Giavarina, "Understanding bland altman analysis," *Biochemi. Medica*, vol. 25, no. 2, pp. 141–151, 2015.

[61] J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, vol. 327, no. 8476, pp. 307–310, Feb. 1986.

[62] O. Kiss *et al.*, "Detailed heart rate variability analysis in athletes," *Clin. Auto. Res.*, vol. 26, no. 4, pp. 245–252, 2016.

DIANA P. TOBÓN received the B.Sc. degree in electronic engineering and the M.Sc. degree in engineering with a subject on the quality of service in wireless body area networks (WBAN) from the University of Antioquia, Medellín, Colombia, in 2008 and 2012, respectively, and the Ph.D. degree from INRS-EMT, Montreal, Canada, in 2017. She continued as a Post-Doctoral Fellow for one term involving on machine learning tools applied in neuromarketing at INRS-EMT. Since 2017, she has been a Post-Doctoral Fellow with the Multimedia Communications Research Laboratory, University of Ottawa, Canada. Her research interests include WBANs, signal processing of noisy biomedical signals, machine learning, and deep learning.

SRINIVASAN JAYARAMAN received the bachelors' degree in electronics and instrumentation engineering from Bharathidasan University, India, the M.Tech. degree in biomedical engineering from SASTRA University, Thanjavur, India, and the Ph.D. degree from the School of Bioscience, IIT Madras, Chennai, India. From 2008 to 2013, he was a Scientist and leading the embedded biomedical Group, TCS Innovation Labs. From 2014 to 2016, he was a Scientist with TCS Innovation Labs, Bangalore, leading the Complementary and Alternative Medicine Research Team. In 2013, he was a Scientific Manager with Trauma Mechanical Labs, University of Nebraska Lincoln, USA, and the New Jersey Institute of Technology, USA. In 2016, he joined as a Post-Doctoral Fellow with INRS-EMT, Canada. He is currently a Research Associate with the Desautels Faculty of Management, McGill, Canada. His research interests include human behavioral and performance modeling, ontology, ergonomics, personalized diagnosis systems, wearable devices, biosignal processing, and human-machine interfaces.

TIAGO H. FALK (SM'14) received the B.Sc. degree from the Federal University of Pernambuco, Brazil, in 2002, and the M.Sc. and Ph.D. degrees from Queen's University, Canada, in 2005 and 2008, respectively, all in electrical engineering. In 2010, he joined INRS, Montreal, Canada, where he is currently an Associate Professor and heads the Multimedia/Multimodal Signal Analysis and Enhancement Laboratory. His research interests include multimedia/biomedical signal analysis and enhancement, pattern recognition, and their interplay in the development of biologically inspired technologies.